# Computerized 'Rosetta Stone' reconstructs ancient languages

February 11 2013

University of British Columbia and Berkeley researchers have used a sophisticated new computer system to quickly reconstruct protolanguages – the rudimentary ancient tongues from which modern languages evolved.

The results, which are 85 per cent accurate when compared to the painstaking manual reconstructions performed by linguists, will be published next week in the Proceedings of the National Academy of Sciences.

"We're hopeful our tool will revolutionize historical linguistics much the same way that statistical analysis and computer power revolutionized the study of evolutionary biology," says UBC Assistant Prof. of Statistics Alexandre Bouchard-Côté, lead author of the study.

"And while our system won't replace the nuanced work of skilled linguists, it could prove valuable by enabling them to increase the number of modern languages they use as the basis for their reconstructions."

Protolanguages are reconstructed by grouping words with common meanings from related modern languages, analyzing common features, and then applying sound-change rules and other criteria to derive the common parent.

The new tool designed by Bouchard-Côté and colleagues at the

University of California, Berkeley analyzes sound changes at the level of basic phonetic units, and can operate at much greater scale than previous computerized tools.

The researchers reconstructed a set of protolanguages from a database of more than 142,000 word forms from 637 Austronesian languages—spoken in Southeast Asia, the Pacific and parts of continental Asia.

  **More information:** Most protolanguages do not leave written records—but in some instances reconstructions can be partially verified against ancient texts or literary histories. A notable exception is well-documented Latin, the protolanguage of the Romance languages, which include modern French, Italian, Portuguese, Romanian, Catalan and Spanish.

Provided by University of British Columbia